

Abstract

A data collection system includes a plurality of node host computers. Each node host computer has a node manager, at least one processing node and a data manager. The data manager delivers network records between processing nodes. The processing nodes include an input queue and an output queue. In some embodiments at least one of the nodes has a plurality of output queues. The processing nodes are arranged in configurable chains of processing nodes. The chains are disposed across one or more of the plurality of host computers to process network records. The at least one processing node having the plurality of output queues can couple records from the plurality of output queues to input queues of corresponding processing nodes, based on a selected distribution method and/or selected content of records. The system can include an administrative client that displays a graphical user interface and an administrative server communicating with the node host computers and the administrative client. The administrative server stores files of executable code that can be dynamically downloaded from the server to configure the client and change the graphical user interface dynamically while the graphical user interface is running. The system includes a order enhancer node that can remove duplicate records produced from gathering statistics concerning network data packets and can order records for delivery to subsequent nodes in the system.